Verification-Based Model Tuning

Caren Marzban
Applied Physics Laboratory, University of Washington
1013 N.E. 40th St.
Seattle, WA 98105-6698
phone: (206) 221-4361 fax: (206) 545-4404 email: marzban@stat.washington.edu

David W. Jones
Applied Physics Laboratory, University of Washington
1013 N.E. 40th St.
Seattle, WA 98105-6698
phone: (206) 543-3236 fax: (206) 545-4404 email: dwjones@apl.washington.edu

Scott A. Sandgathe
35574 Camp Creek Rd.
Springfield, OR 97478
phone: (541) 988-0289 fax: (541) 988-0289 email: sandgathe@apl.washington.edu

Award Number: N00014-12-1-0276

LONG-TERM GOALS

All numerical models (e.g., Numerical Weather Prediction models) have certain parameters within model algorithms which effect forecasts to a different degree, depending on the forecast quantity. The specific values of these model parameters are determined either theoretically using fundamental physics laws but incorporating necessary approximations to reduce computational cost, or empirically using observations from field experiments where observational error introduces uncertainty. In either case, the exact value of the parameter is often unknown *a priori*, and so their values are usually set to improve forecast quality through some form of forecast verification. Such an approach to model tuning, however, requires knowledge of the observations to which the forecasts must be compared, and therefore, a multitude of highly detailed experimental cases in order to fully resolve parameter values, a data set very difficult to obtain. A knowledge of the relationship between model parameters and forecast quantities, without reference to observations, can not only aid in such an observation-based approach to model tuning, it can also aid in tuning the model parameters according to other criteria that may not be based on observations directly, e.g., a desire to affect the forecasts according to some long-term experience of a forecaster. The main goal of our work has been to develop a framework for representing the complex relationship between model parameters and forecast quantities, without any reference to observations.

OBJECTIVES

The specific objective has been to develop and test the aforementioned framework, first on the Lorenz '63 model, and then on COAMPS.

Report Documentation Page	*Form Approved* *OMB No. 0704-0188*

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
30 SEP 2013		**00-00-2013 to 00-00-2013**

4. TITLE AND SUBTITLE	
Verification-Based Model Tuning	5a. CONTRACT NUMBER
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **University of Washington,Applied Physics Laboratory,1013 N.E. 40th St.,Seattle,WA,98105**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**8**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

APPROACH

There are numerous ways of representing the relationship between a set of input variables (e.g., model parameters) and a set of output variables (e.g., forecast quantities). A most common way is to simply develop a multivariate regression model relating the inputs to the outputs. However, such an approach presumes that 1) the choice of the inputs, and their interactions, are already known; it also 2) requires specifying which varaiables should be inputs and which should be outputs, because in regression the inputs and outputs are not treated symmetrically. In practice, however, it is often not known which inputs affect the outputs, and which do not. Also, it is often desirable to represent the relationship between the inputs and the outputs in a symmetric way, similar to the notion of a correlation. We address the first point by using a Variance-Based Sensitivity Analysis (VBSA), which is capable of assessing the predictive strength of each input variable (e.g., model parameter), the interaction between them, and their statistical uncertainty. The symmetric assessment of the association between the inputs and the outputs is performed through a Canonical Correlation Analysis (CCA).

It is important to point out that all of the above approaches require only model forecasts, and no analysis or observations. This is an important feature of the method because model forecasts are generally and significantly easier to generate than observations.

WORK COMPLETED

The VBSA has now been applied to the Lorenz '63 model, and the results have been compared to an adjoint-based approach to sensitivity analysis (Marzban 2013).

VBSA has also been applied to COAMPS (Marzban, Sandgathe, Doyle, Lederer 2013).

CCA has been employed to determine the combination of model parameters, and the combination of forecast quantities, which are most correlated with one another (Marzban, Sandgathe Doyle 2013)..

RESULTS

The results of the Lorenz '63 model are twofold: First, we demonstrated that the VBSA can be applied to a reasonably complex, nonlinear model. More substantively, we showed that the Z state variable in the Lorenz model (normally representing vertical temperature variation) is most sensitive to the r and b parameters (related to the Raleigh number and wavenumber, respectively). We also found there is evidence for an interaction between these two model parameters. Finally, we showed that the VBSA method produces similar results to the adjoint-based sensitivity analysis method, with the exception that the VBSA additionally provides measures of uncertainty. An example of all of the results is shown in Figure 1, where the distributions of some sensitivity measures (V, VT, S, ST), and measures of interaction, are shown. The boxplots summarize the underlying distributions in a way that displays their uncertainty. Also, we find that Latin Hypercube sampling of the parameter space (grey boxplots) produces more precise estimates than simple random sampling.

The VBSA method was applied to eleven model parameters in COAMPS and four forecast quantities. A complete list of the parameters is shown in Table 1; the four forecast quantities are 24hr forecasts of convective, stable, and total precipitation, and accumulated snow. Regarding convective precipitation, we find the most influential parameter to be the fraction of available precipitation in the Kain-Fritsch cumulus parameterization fed back to the grid scale. Stable and total precipitation are most affected by

a linear factor that multiplies the surface fluxes; and the parameter that most affects accumulated snow is the microphysics slope intercept parameter for snow. Furthermore, all of the interactions between the parameters are found to be either exceedingly small, or have too much variability (across days and/or parameter values) to be of primary concern Figure 2 and 3 provide two visual displays of all the sensitivities and interactions for convective precipitation. Although it is natural for the eye to be drawn to the parameters (and interactions) with the highest sensitivities, it is important to point out that knoweldge of the parameters with near-zero sensitivity is also important because one can them neglect them in model tuning

Although not shown in this report, the above VBSA showed that the four forecast quantities are controlled by different sets of model parameters. This finding implies that optimally setting the model parameters in a way to affect one forecast quantity may have adverse effects on other forecast quantities. As such, it is natural to ask what combination of the forecast quantities is optimally controlled by the model parameters. CCA is designed to answer a similar question: What linear combination of the forecast quantities, and what linear combination of the model parameters, are maximally correlated? These linear combinations are called Canonical Variates (CV). We showed that the forecast quantities are indeed correlated, and that their CVs are controlled by specific model parameter CVs. The coefficients in the linear combinations, also called loadings, measure the degree to which model parameters contribute to their CV, and the degree to which a forecast quantity contributes to its CV. Figure 4 shows these loadings for COAMPS. The symbols denoting the parameters are shown in Marzban, Sandgathe, and Doyle (2013); conv, stab, and snow denote convective precipitation, stable precipitation, and accumulated snow. The boxplots display both daily variability (across 36 days sampled between January and July, 2009) and sampling variability in parameter space. From these results it follows that the model parameters can be set to affect the sum, and the difference between convective and stable precipitation, while keeping snow mostly constant; a different combination of model parameters is shown to mostly affect the difference between stable precipitation and snow, with minimal effect on convective precipitation.

IMPACT/APPLICATIONS

All of these results can be used to better set model parameters for the purpose of improving forecasts.

RELATED PROJECTS

There are no related projects, but an NSF proposal is in preparation wherein the forecast quantities are **spatial features** of the forecasts, e.g., the number of “objects,” their location, and orientation (as approximated by the axes of an ellipse) in a forecast field.

PUBLICATIONS

Marzban, C. 2013: Variance-based sensitivity analysis: An illustration on the Lorenz '63 model. *Monthly Weather Review*. [in press, refereed]

Marzban, C., Scott Sandgathe, James D. Doyle 2013: Model tuning with canonical correlation analysis. *Monthly Weather Review*. [conditionally accepted, refereed]

Marzban, C., Scott Sandgathe, James D. Doyle, Nicholas C. Lederer 2013: Variance-based sensitivity analysis: Preliminary results in COAMPS. *Monthly Weather Review*. [conditionally accepted, refereed]

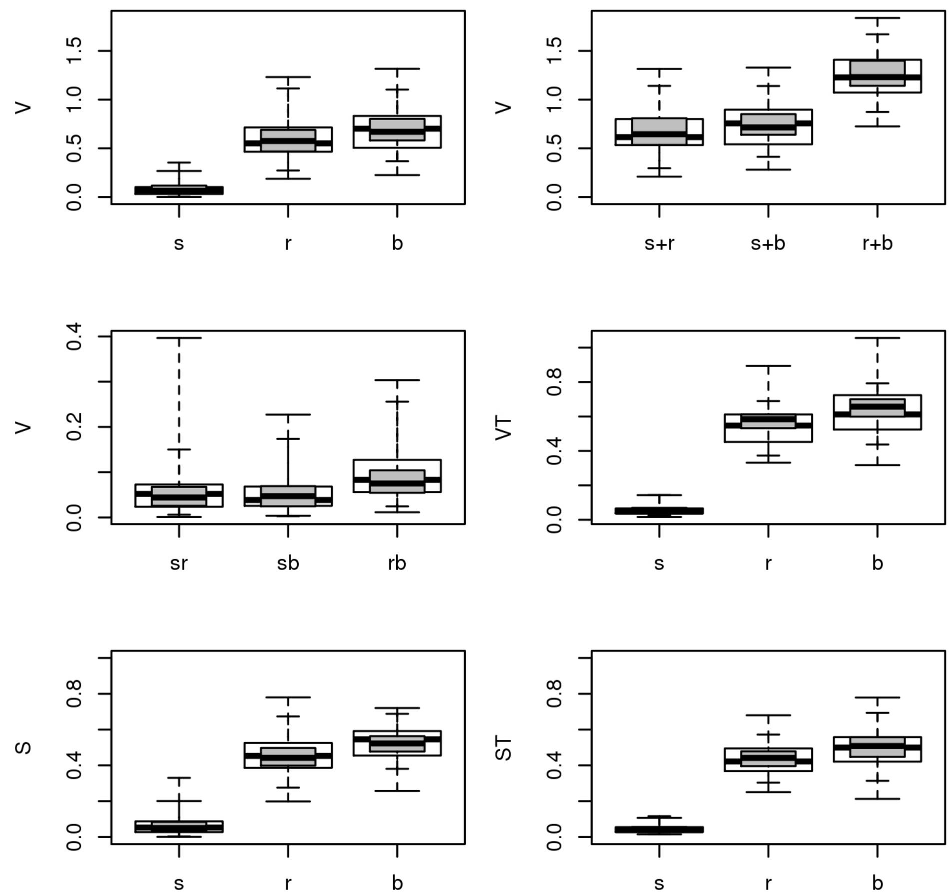

Figure 1. The distribution of sensitivity and interaction measures for the Lorenz '63 model. The white boxplots correspond to a simple random sampling of the model parameter space, and the grey boxplots correspond to a latin hypercube sampling. It can be seen that the latter provide more precise estimates.

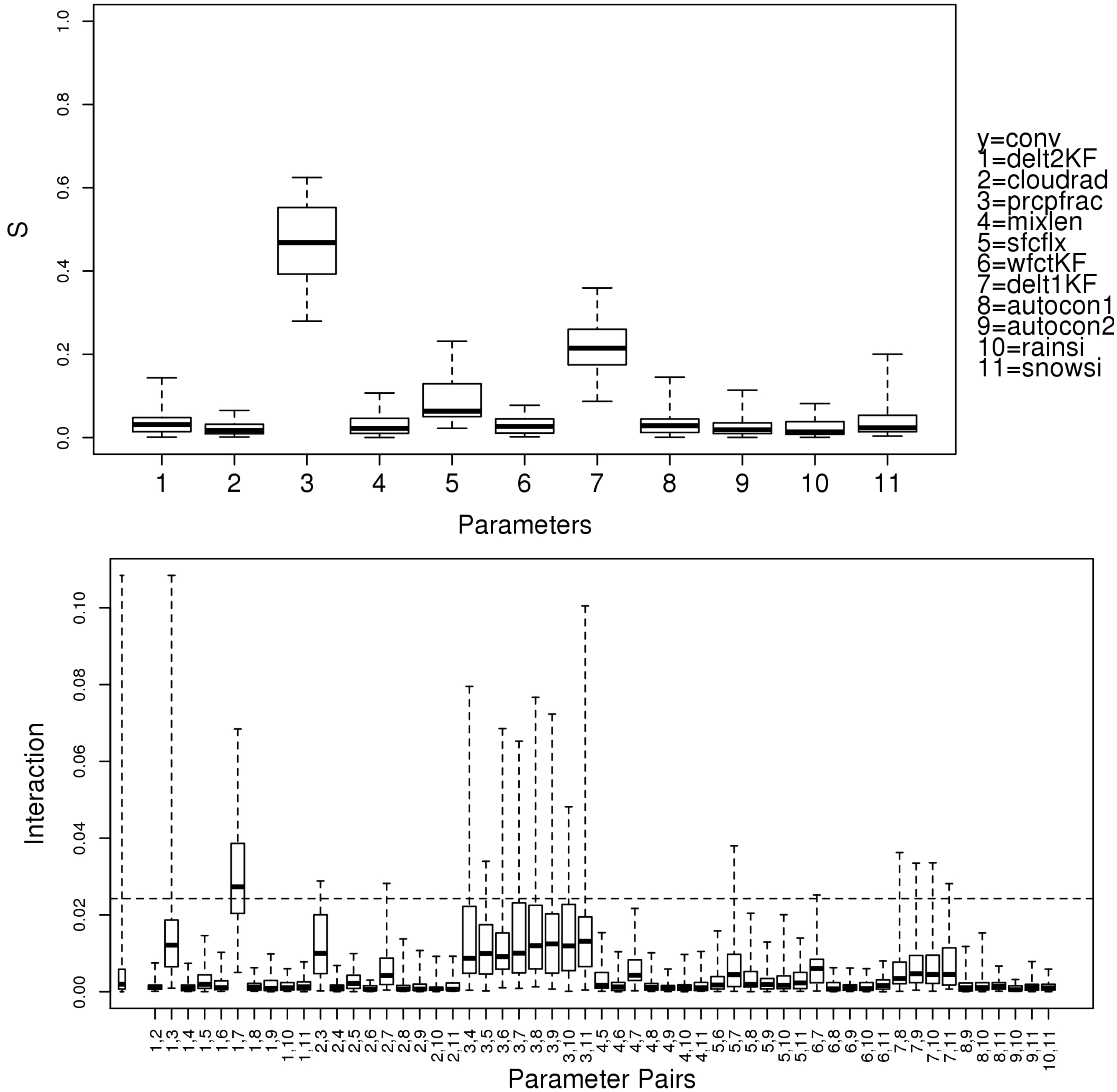

Figure 2: Distribution/boxplot of sensitivity of convective precipitation on model parameters (top) and their interactions (bottom). The model parameters are listed in the top/right corner, and their detailed description can be found in Table 1.

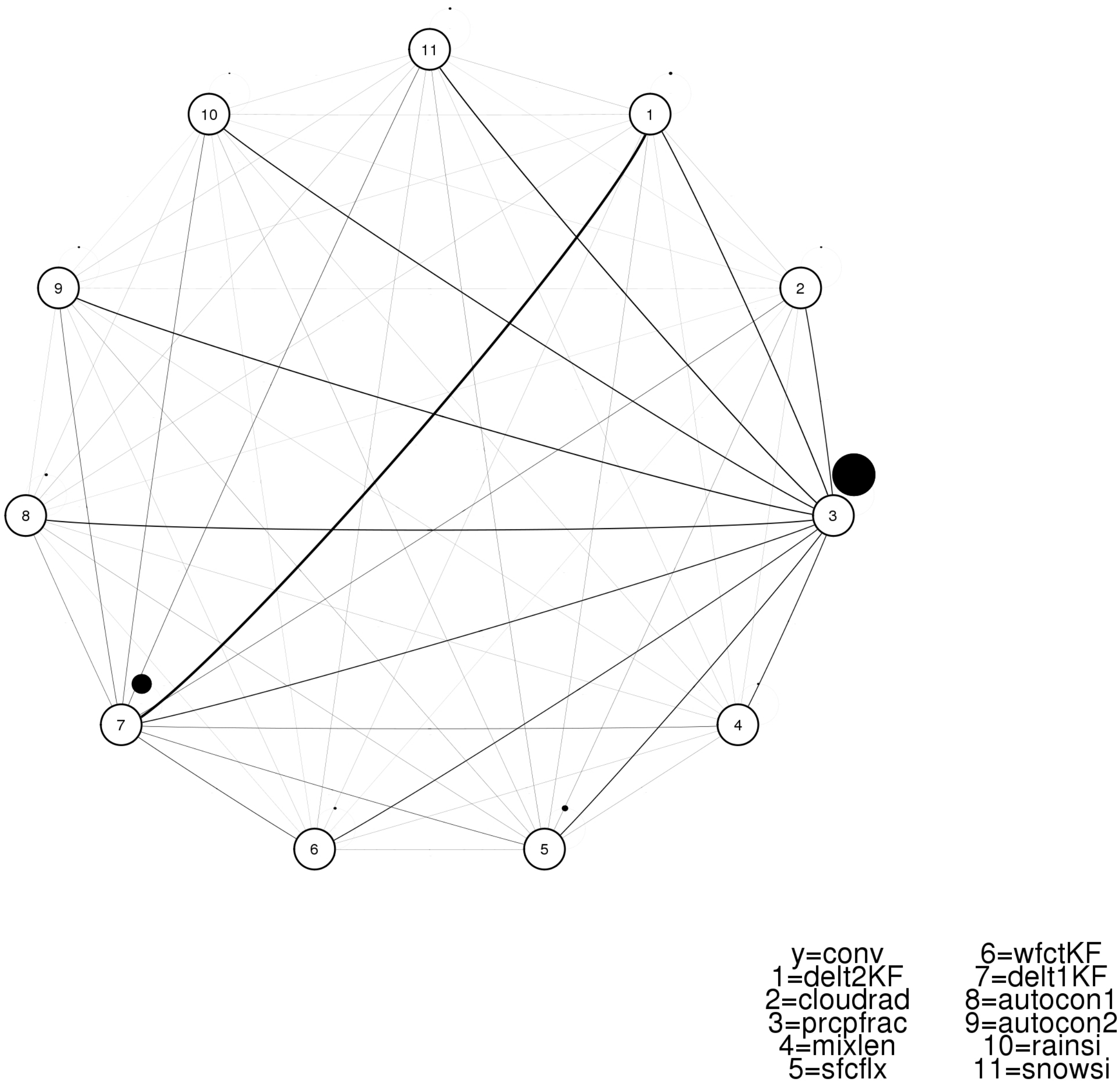

Figure 3: A network diagram for visualizing the sensitivities and the interaction between the model parameters.

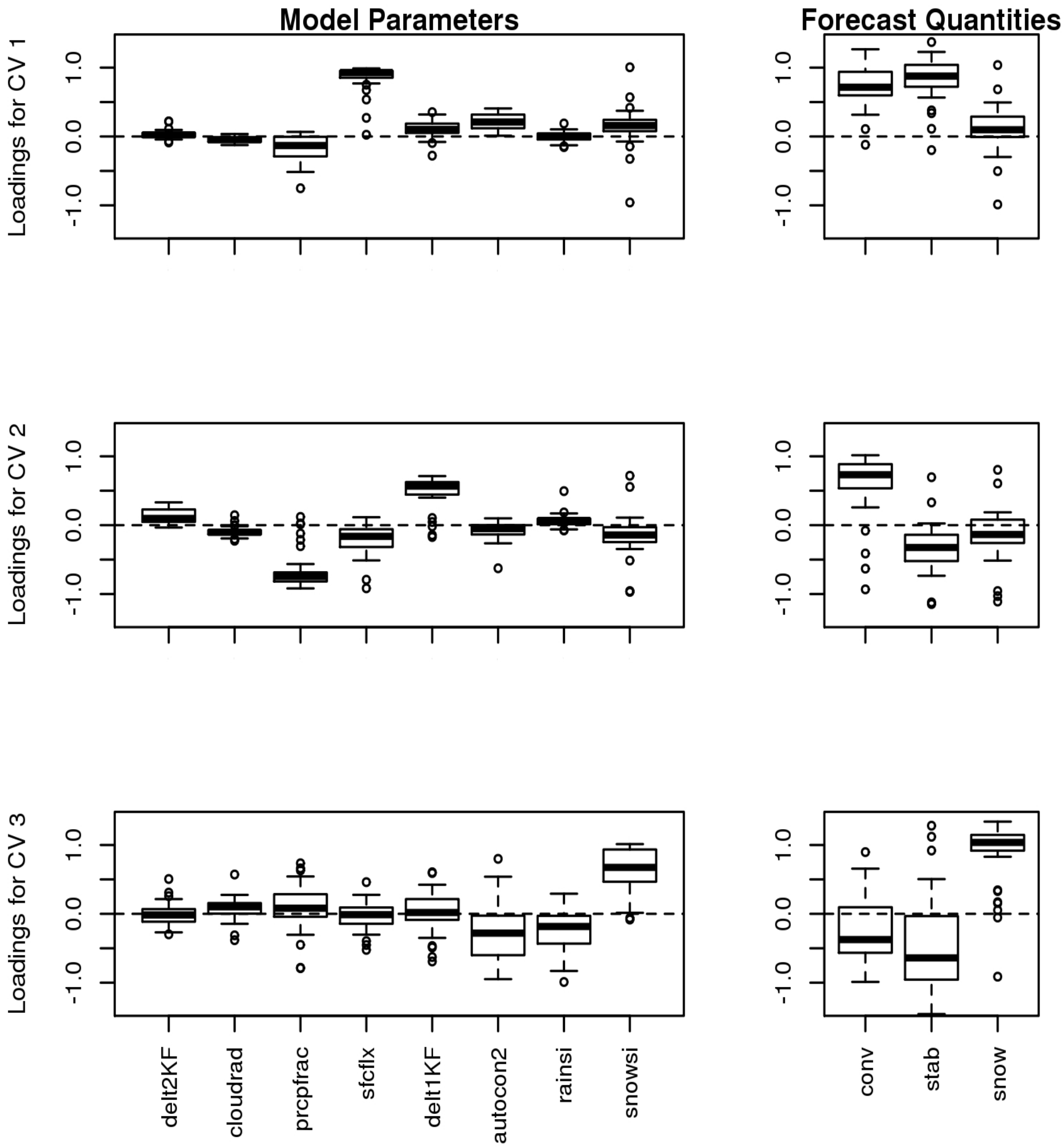

Figure 4. The loading of the model parameters (left), and the loading for the forecast quantities (right) when the corresponding CVs are most correlated (top), i.e. correlation coefficient ~ 0.95. The two loadings on two lower-correlated CVs (correlation coefficient ~ 0.85, and 0.70) are also shown (middle and bottom rows).

Table 1. The COAMPS model parameters, their default values, and the range used in the study.

Name (Unit)	Description	Default	Range
mixlen	Linear factor that multiplies the mixing length within the PBL	1.0	0.5, 1.5
sfcflx	Linear factor that modifies the surface fluxes	1.0	0.5, 1.5
wfctKF	Linear factor for the vertical velocity (grid scale) used by KF trigger	1.0	0.5, 1.5
delt1KF ($^\circ C$)	Temperature increment at the LCL for KF trigger	0	-2, 2
delt2KF ($^\circ C$)	Another method to perturb the temperature at the LCL in KF	0	-2, 2
prcpfrac	Fraction of available precipitation in KF, fed back to the grid scale	0.5	0, 1
cloudrad (m)	Cloud radius factor in KF	1500	500, 300
autocon1 ($\frac{kg}{m^3 s}$)	Autoconversion factors for the microphysics	0.001	1e-4, 1e-
autocon2 ($\frac{kg}{m^3 s}$)	Autoconversion factors for the microphysics	4e-4	4e-5, 4e-
rainsi ($\frac{1}{m}$)	Slope intercept parameter for rain in the microphysics	8.0e6	8.0e5, 8.
snowsi ($\frac{1}{m}$)	Slope intercept parameter for snow in the microphysics	2.0e7	2.0e6, 2.

KF = Kain-Fritsch

PBL = Planetary Boundary Layer

LCL = Lifted Condensation Level